

Обучение машинного перевода без параллельных текстов

Борисова А. В. Бажтеев¹ О. Ю. Стрижов² В. В.

¹Московский физико-технический институт

²Вычислительный центр им. А. А. Дородницына ФИЦ ИУ РАН

В данной работе рассматривается задача построения системы машинного перевода. Для построения систем машинного перевода используются параллельные предложения. Однако не для всех пар языков существует достаточный массив размеченных предложений. В данной работе рассматривается метод построения модели машинного перевода без использования параллельного корпуса. Модель основана на Seq2Seq, предложениям обоих языков ставится в соответствие вектор в общем скрытом пространстве. Вычислительный эксперимент проводится на паре языков "русский-украинский".

Ключевые слова: *машинный перевод, Seq2Seq.*